TURING'S THINKING MACHINE AND 'T HOOFT'S PRINCIPLE OF SUPERPOSITION OF STATES

ZEQIAN CHEN

Abstract. In his 1950 paper [11], Turing proposed the notion of a thinking machine, namely a machine that can think. But a thinking machine has to follow a certain principle of physics, provided it is realized physically. In this paper, we show that Turing's thinking machine necessarily obeys 't Hooft's principle of superposition of states, which was presented by 't Hooft [8] in 2016 beyond the usual one as described by Dirac [4] in the conventional quantum mechanics. Precisely, Turing's thinking machine must be a quantum machine, while 't Hooft's principle characterizes its thinking behavior in a probabilistic way.

In his 1950 paper [11], Turing considered the question: Can machines think? As noted by him, this should begin with definitions of the meaning of the terms "machine" and "think". He pointed out that a machine that can think is mathematically defined to be a Turing machine. However, instead of giving a definition for thinking, he described the problem in terms of a game named by "imitation game", that is now called "the Turing test". Furthermore, Turing pointed out that a thinking machine must learn knowledge before it can think, as Human intelligence, to no small extent, relies on knowledge.

The Turing machine can be realized physically in many different ways. A physically realized machine has to follow a certain principle of physics, but what is the basic principle that Turing's thinking machine has to follow? In what follows, we show that Turing's thinking machine necessarily obeys 't Hooft's principle of superposition of states, which was presented by 't Hooft [8, Section 2.1] in 2016, asserting as:

't Hooft's principle of superposition of states A quantum system is associated with a Hilbert space $\mathbb{H}$ and consists of physical states $|A\rangle$ which form a basis of $\mathbb{H}$; a quantum state of the form

$$|\psi\rangle = \sum_A \lambda_A |A\rangle \tag{1}$$

describes a situation where the probability to find the system to be in the physical state $|A\rangle$ is $|\lambda_A|^2$.

Note that 't Hooft's principle of superposition of states is different from the usual one as described by Dirac [4] in the conventional quantum mechanics. t' Hooft proposed this principle for resolving the measurement problem of quantum mechanics. Here, we involve this principle for understanding Turing's thinking machine.

For illustrating how 't Hooft's principle of superposition of states is applied to Turing's thinking machine, we give the following example:

$$\begin{tikzcd}
 & & & C \arrow[r, "\text{Output}", "p=0.306"'] & OAC \\
 & & A \arrow[ur, "p=0.51"] \arrow[r, "p=0.49"] & D \arrow[r, "\text{Output}", "p=0.294"'] & OAD \\
\text{Input} \arrow[r] & O \arrow[ur, "p=\frac{3}{5}"] \arrow[dr, "p=\frac{2}{5}"'] & & & \\
 & & B \arrow[r, "p=0.95"] \arrow[dr, "p=0.05"'] & E \arrow[r, "\text{Output}", "p=0.38"'] & OBE \\
 & & & F \arrow[r, "\text{Output}", "p=0.02"'] & OBF
\end{tikzcd}$$

In the case of transformation from O to A or B, there are two ways: One is the classical evolution

$$|O\rangle\langle O| \longrightarrow \frac{3}{5}|A\rangle\langle A| + \frac{2}{5}|B\rangle\langle B|,$$

the other is the quantum evolution

$$|O\rangle \longrightarrow \sqrt{\frac{3}{5}}|A\rangle + \sqrt{\frac{2}{5}}|B\rangle;$$

similarly for the cases from A to C or D and from B to E or F. We remark that

(1) the probabilities $p_{OA} = 0.6, p_{OB} = 0.4$, *et al.*, are depending only on knowledge that the machine has learned, that is, these probabilities represent the knowledge that the machine possesses;
(2) if the machine follows classical physical law, the outcome must be OAC;
(3) if the machine follows quantum physical law, that is 't Hooft's principle of superposition of states, the outcome must be one of four cases OAC, OAD, OBE, and OBF in a probabilistic way according to the probabilities as being $p_{OAC} = 0.306, p_{OAD} = 0.294, p_{OBE} = 0.38$ or $p_{OBF} = 0.02$ respectively;
(4) Turing's thinking machine needs to be a quantum machine, since the outcome OBE can only appear in a quantum way, whose probability $p_{OBE} = 0.38$ is larger than the one $p_{OAC} = 0.306$ of OAC, that is regarded as the result that the machine can think;
(5) Turing's thinking machine needs not to be identical to a quantum computer, because quantum computation is necessarily based on the many-body superposition states (cf.[2]).

This example is a 3-layer neural network, generic neural networks (cf.[1, 10]) can be considered in the same way.

In conclusion, artificial general intelligence should be based on Turing's thinking machine, which has to be a quantum machine (cf.[6]), while 't Hooft's principle of superposition of states characterizes the thinking behavior of an intelligence machine in a probabilistic way. A mathematical formalism for Turing's thinking machine was presented in [3], by utilizing the topos-theoretic formulation of physics developed by Döring and Isham [5].

Appendix: Mathematical formulation of 't Hooft's quantum mechanics Since 't Hooft's principle of superposition of states is distinct from the usual one, the Dirac-von Neumann formalism of quantum mechanics (cf.[9]) needs to modify. That is, the mathematical formalism of 't Hooft's quantum mechanics is defined by a set of postulates as follows:

(P_1) **The state postulate** A quantum system is associated with a Hilbert space $\mathbb{H}$ and consists of physical states $|\psi_k\rangle$ which form a basis of $\mathbb{H}$, called the ontological basis; a quantum state of the form

$$|\psi\rangle = \sum_k a_k|\psi_k\rangle$$

describes a situation where the probability to find the system to be in the physical state $|\psi_k\rangle$ is $|a_k|^2$, provided $a_k \neq 0$ for k.

(P_2) **The observable postulate** An observable for the quantum system is represented by a Hermitian operator in $\mathbb{H}$, which is diagonal in the ontological basis and called an ontological observable or a beable; if $B = \sum_k \lambda_k|\psi_k\rangle\langle\psi_k|$ is an ontological observable, then the expectation of B at a quantum state $|\psi\rangle = \sum_k a_k|\psi_k\rangle$ is given by

$$\langle B\rangle_\psi = \langle\psi, B\psi\rangle; \tag{2}$$

in particular, the probability of obtaining the outcome λ_k is $|a_k|^2$ for any k.

(P_3) **The evolution postulate** The evolution of the system is described by a family $\{U(t) : t \in \mathbb{R}\}$ of unitary operators, satisfying that the mapping $t \to U(t)$ is strongly continuous, $U(0) = I$, and

$$U(t)U(s) = U(t+s), \quad \forall s, t \in \mathbb{R}, \tag{3}$$

and such that if the initial state is the physical state ψ_k, then the system is at the state $U(t)\psi_k$ for any time t.

(P_4) **The composite-systems postulate** The Hilbert space associated with a composite quantum system is the Hilbert space tensor product of the Hilbert spaces of its components. If systems numbered 1 through n are prepared in physical states $\psi^{(j)}$, $j = 1, \ldots, n$, then the joint physical state of the composite total system is the tensor product $\psi^{(1)} \otimes \cdots \otimes \psi^{(n)}$.

For illustration, we consider the Deutsch-Jozsa quantum algorithm (cf.[7]). Given a function $f : \{0,1\}^n \mapsto \{0,1\}$, define $U_f : (\mathbb{C}^2)^{\otimes n} \otimes \mathbb{C}^2 \mapsto (\mathbb{C}^2)^{\otimes n} \otimes \mathbb{C}^2$by

$$U_f|\vec{x}, y\rangle = |\vec{x}, f(\vec{x}) \oplus y\rangle, \quad \forall \vec{x} \in \{0,1\}^n,\ y \in \{0,1\},$$

where $a \oplus b = (a+b)\text{mod}2$. Choose $\psi_0 = |0\rangle^{\otimes n} \otimes |1\rangle$ as an initial state and $B = \sigma_z \underbrace{\otimes \cdots \otimes}_{n+1} \sigma_z$ to be the ontological observable. Let $U = U_3 U_2 U_1$, where

$$U_1 = H^{\otimes(n+1)},\ U_2 = U_f,\ U_3 = H^{\otimes n} \otimes I_2,$$

and H is the Hadamard transformation on $\mathbb{C}^2$. By computation, one has,

$$\phi = U_3 U_2 U_1 |0\rangle^{\otimes n} \otimes |1\rangle = \sum_{\vec{x}, \vec{y} \in \{0,1\}^n} \frac{(-1)^{\vec{x}\cdot\vec{y} + f(\vec{x})}|\vec{y}\rangle}{2^n} \otimes \frac{|0\rangle - |1\rangle}{\sqrt{2}},$$

see [7] for the details. Then, we make the measurement of B at ϕ and obtain the required results: Since the coefficients of the first n-qubit state $|0\rangle^{\otimes n}$ is $\sum_{\vec{x}} (-1)^{f(\vec{x})}/2^n$, if f is a constant function, the absolute value of this coefficient is 1, so all the outcomes of measurement on σ_z at the first n-qubit are 1; or if f is balanced, the coefficient is zero, so there are some of the outcomes which must be -1.

References

[1] K. Beer, D. Bondarenko, T. Farrelly, T.J. Osborne, R. Salzmann, D. Scheiermann, R. Wolf, Training deep quantum neural networks, *Nature Communications* **11** (2020), 808: 1-6.

[2] Z. Chen, Observable-geometric phases and quantum computation, *International Journal of Theoretical Physics* **59** (2020), 1255-1276.

[3] Z. Chen, L. Ding, H. Liu, J. Yu, Topos-theoretic formalism of quantum artificial intellegence (in Chinese), preprint, 2024.

[4] P.A.M. Dirac, *The Principles of Quantum Mechanics* (Fourth Edition), Oxford University Press, London, 1958.

[5] A. Döring, C.J. Isham, "What is a thing?": Topos theory in the foundations of physics, *Lecture Notes in Physics* **813**, pp.753-937, Springer-Verlag, Berlin, 2011.

[6] V. Dunjko, H.J. Briegel, Machine learning & artifcial intelligence in the quantum domain: a review of recent progress, *Reports on Progress in Physics* **81** (2018), 074001: 1-68.

[7] S. Gudder, Quantum computation, *The American mathematical monthly* **110** (2003), 181-201.

[8] G. 't Hooft, *The Cellular Automaton Interpretation of Quantum Mechanics,* Springer, 2016.

[9] J. von Neumann, *Mathematical Foundations of Quantum Mechanics,* Princeton University Press, Princeton, 1955.

[10] K. Sharma, M. Cerezo, L. Cincio, P.J. Coles, Trainability of dissipative perceptron-based quantum neural networks, *Physical Review Letters* **128** (2022), 180505: 1-7.

[11] A. Turing, Computing machinery and intelligence, *Mind* **59** (1950), 433-460.

Wuhan Institute of Physics and Mathematics, Innovation Academy for Precision Measurement Science and Technology, Chinese Academy of Sciences, 30 West District, Xiao-Hong-Shan, Wuhan 430071, China.